J Digit Imaging (2013) 26:173–182
DOI 10.1007/s10278-012-9497-z

Joint Segmentation and Groupwise Registration of Cardiac Perfusion Images Using Temporal Information

Dwarikanath Mahapatra

Published online: 12 June 2012

Abstract We propose a joint segmentation and groupwise registration method for dynamic cardiac perfusion images that uses temporal information. The nature of perfusion images makes groupwise registration especially attractive as the temporal information from the entire image sequence can be used. Registration aims to maximize the smoothness of the intensity signal while segmentation minimizes a pixel's dissimilarity with other pixels having the same segmentation label. The cost function is optimized in an iterative fashion using B-splines. Tests on real patient datasets show that compared with two other methods, our method shows lower registration error and higher segmentation accuracy. This is attributed to the use of temporal information for groupwise registration and mutual complementary registration and segmentation information in one framework while other methods solve the two problems separately.

Keywords Groupwise registration · Segmentation · Temporal information · Cardiac · Perfusion · MRI

Introduction

Dynamic contrast-enhanced (DCE) magnetic resonance (MR) images (or perfusion MRI) has developed as a popular non-invasive tool for the functional analysis of internal organs. Contrast agent is injected intravenously into the patient and a series of MR scans are acquired over a period of time. As the contrast agent flows through the blood stream, the intensity of those regions increases. Since the image acquisition process can take up to 20 min, patient movement is inevitable. Although the cardiac images are acquired by electrocardiogram gating, there are some residual deformations due to patient breathing which needs correction. Perfusion images are characterized by rapid intensity change over time, low spatial resolution and noise, and make registration challenging. Previous techniques mostly employed a pairwise registration approach, i.e., all images of a sequence are individually registered to a fixed reference image [1, 2]. The success of such approaches depends upon the robustness of the cost function to intensity change. Although intensity change due to contrast agent flow makes registration challenging, it also provides important temporal information for registration and segmentation of the perfusion image sequence. In this work, we propose a method which makes use of the temporal dynamics of contrast agent flow to achieve joint segmentation and groupwise registration of an MR cardiac image sequence.

The changing intensity due to contrast agent flow provides valuable segmentation information by highlighting the organ of interest. It is interesting to note that different regions of a cardiac image sequence have different intensity time characteristics. Temporal flow information was used for registration [3] and segmentation [4] of perfusion images. In [5] the registration, framework constrained the deformation field such that different regions follow a particular intensity time profile. It is an accepted fact that improved registration leads to accurate segmentation and vice versa. In this paper, we make use of temporal information to achieve registration and segmentation of the entire image sequence. Instead of pairwise registration, we solve the problem using groupwise registration as it allows us to impose constraints based on temporal information.

D. Mahapatra (✉)
Department of Computer Science,
Swiss Federal Institute of Technology (ETH) Zurich,
Room CAB F 61.1, Universitätstrasse 6,
8092 Zurich, Switzerland
e-mail: dwarikanath.mahapatra@inf.ethz.ch

Of late, groupwise registration methods have gained popularity because of a need to register large number of datasets for atlas construction [6, 7]. Groupwise registration is generally approached by two techniques. The first approach uses pairwise registration between a template and all other images in the population as in [8–10]. Yang et al. in Ref. [10] utilize the voxel-wise geometry and orientation information for registration. Rotation invariant features are used in a deformable matching mechanism for registration. However, pairwise registration has two limitations. First, selecting a fixed template may not accurately represent the population. Secondly, pairwise registration is not very effective when registering two images with significant anatomical differences. Only those subjects that are close to the template are registered properly.

To overcome the above limitations, there are many methods that achieve registration using all images from the population. This approach is more faithful to the term "groupwise registration." The goal is to simultaneously warp all subjects in a population towards a hidden common space [11, 12]. The groupwise registration problem is formulated as one of optimization, with a global cost function defined on all aligned images [13, 14]. The cost function in Refs. [13, 14] is the stack entropy or the entropy of corresponding voxels in different volumes. If volumes are properly aligned intensity values at corresponding voxels from all volumes will form a low entropy distribution. With this constraint, groupwise registration is achieved within a B-spline-based freeform deformation framework.

When across subject variation is very large, it is generally difficult to achieve good registration by simply registering each image to a template image. Therefore many approaches make use of intermediate templates for registration [15–18]. In Refs. [15, 18], an intermediate template which does not belong to the original dataset is created to aid registration between two images. Tang et al. in Ref. [18] warp the template image with a set of simulated deformation fields learned using principal component analysis (PCA) on a set of training deformation fields. Such approaches do not guarantee that the intermediate template is realistic which may affect registration results. Jia et al. in Ref. [16] construct a tree structure where each node of the tree is represented by an image, and similar images are represented by connected nodes. Each image is registered with the help of intermediate templates determined along its own path with respect to the final template.

Groupwise registration methods are particularly suitable for perfusion DCEMRI. Each region of the scanned organ is characterized by a different intensity profile over time. For example, in cardiac perfusion MRI the contrast agent first flows into the right ventricle (RV) and then into the left ventricle (LV) before being flushed out of the cardiovascular system. Thus, a pixel within the RV shows a peak intensity magnitude early in the scanning sequence while for the LV blood pool, the intensity peak occurs later (Fig. 3). With this available information, we can formulate the cost function such that after registration, pixels from certain regions follow a particular intensity time profile. This is achieved by joint segmentation and groupwise registration of DCE-MRI. There are not many works dealing with joint segmentation and groupwise registration of cardiac perfusion MRI, although Zhang et al. in Ref. [19] describe a method for the rigid registration of brain perfusion images. The cost function is derived from the total quadratic variation of image intensity. Metz et al. in Ref. [20] propose a method for groupwise registration of dynamic lung data using both spatial and temporal constraints where groupwise optimization of B-splines is used.

The temporal intensity patterns of pixels also determine their segmentation labels (i.e., RV, LV blood pool, myocardium background, etc). Thus, the image sequence can be segmented along with groupwise registration. It is a well-known fact that registration and segmentation are mutually complementary approaches. Many works have combined them in a joint registration and segmentation framework [21–23]. In previous works, we have proposed a method for the joint registration and segmentation of natural and cardiac perfusion images [24, 25]. However, these methods used pairwise registration of images without exploiting temporal information. Including segmentation information into the cost function reduces registration error and improves segmentation accuracy. In this paper, we propose a joint segmentation and groupwise registration (JSGR) approach for cardiac perfusion MRI. Our method combines intensity information from the entire image sequence (for registration) and maximizes the similarity between a pixel and other pixels belonging to the same class (for segmentation). We describe our method in "Materials and Methods", present experimental results in "Experiments and Results", and conclude with "Conclusions."

Materials and Methods

JSGR aims to find the transformation for each image that minimizes a cost function. We do not have an explicitly defined reference image for elastic registration. The transformations are constrained such that the registered images approach a common image space which is approximately the mean image of the dataset. This mean image changes with each registration iteration and is the reference image. Wu et al. in Ref. [26] have highlighted the importance of a sharp mean image for accurate groupwise registration. Their observations are derived from constructing atlases for a

large population of brain images. Such datasets show a lot of variability and a fixed reference image is sure to introduce bias. Our method is applied on one patient dataset at a time and hence a sharp mean image at the beginning is not necessary. The general objective function for JSGR consists of two terms, i.e.,

$$E = E_{\text{data}} + E_{\text{smooth}}, \tag{1}$$

where E_{data} is the data cost and E_{smooth} is the smoothness cost. The data cost depends upon the type of images being registered and the smoothness cost depends upon the optimization framework. Perfusion images are characterized by rapid intensity change over time. Instead of relying on low level information we aim to exploit the temporal information for groupwise registration. First, we give a brief description of B-splines and their optimization. Then we explain the formulation of our data cost (E_{data}).

B-spline Based Registration

A B-spline-based freeform deformation (FFD) transformation model was presented in [27] for the elastic registration of breast MRI. The basic idea of FFDs is to deform an object by manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3D (or 2D) object and produces a smooth and continuous transformation. The transformation field consists of a global and local component and is defined as

$$T(x) = T_{\text{local}}\left(T_{\text{global}}(x)\right), \tag{2}$$

where T_{global} is a rigid transform obtained using Ref. [28], and T_{local} is the deformation based on B-splines. The DCE-MR images are first rigidly registered to a chosen reference image. This reference image is only for the purpose of rigid registration and is chosen to be the image from the dynamic sequence that shows all tissues without ambiguity (e.g., the image corresponding to the time point of peak enhancement such that all tissues are clearly shown). For elastic registration, there is no explicitly defined reference image. Further discussion is restricted to T_{local}.

We define an initial $n_x \times n_y \times n_z$ grid of control points denoted as Φ. The grid points are denoted as $\Phi_{i,j,k}$ and have uniform spacing. The freeform deformation can be written as the 3D tensor product of 1D cubic B-splines,

$$T_{\text{local}}(x) = x + \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)\Phi_{i+l,j+m,k+n}, \tag{3}$$

where $x=(x_1, x_2, \text{and } x_3)$ is the displacement vector, $i=\lfloor x=nx \rfloor - 1$, $j=\lfloor y=ny \rfloor - 1$, $k=\lfloor z=nz \rfloor - 1$, $u=x=nx-\lfloor x=nx \rfloor$, $v=y=ny-\lfloor y=ny \rfloor$, $w=z=nz-\lfloor z=nz \rfloor$ and Bl is the lth cubic B-spline basis function given by the following equations:

$$\begin{aligned} B_0(u) &= \tfrac{(1-u)^3}{6} \\ B_1(u) &= \tfrac{3u^3-6u^2+4}{6} \\ B_2(u) &= \tfrac{-3u^3+3u^2+3u+1}{6} \\ B_3(u) &= \tfrac{u^3}{6} \end{aligned} \tag{4}$$

Since our datasets are in 2D, the corresponding equation is

$$T_{\text{local}}(x) = x + \sum_{l=0}^{2}\sum_{m=0}^{2} B_l(u)B_m(v)\Phi_{i+l,j+m}, \tag{5}$$

B-splines are locally controlled, which makes them computationally efficient even for a large number of control points. In particular, the basis functions of cubic B-splines have a limited support, i.e., changing control point $\Phi_{i,j}$ affects the transformation only in the local neighborhood of that control point. The local deformation of the cardiac tissues should be characterized by a smooth transformation. The general form of such a transformation in 2D takes the following form:

$$E_{\text{smooth}} = \frac{1}{A}\int_0^X \int_0^Y \left[\left(\frac{\partial^2 T}{\partial x^2}\right)^2 + \left(\frac{\partial^2 T}{\partial y^2}\right)^2 + 2\left(\frac{\partial^2 T}{\partial xy}\right)^2\right], \tag{6}$$

where A is the image area.

Similarity Measure

The images are first rigidly aligned with respect to a reference image using [29]. Note that this reference image is only for rigid alignment and is not used for groupwise registration. Seed points belonging to RV, LV blood pool, myocardium, and background are identified (as shown by red arrows in the first image of the first row of Fig. 2), and the labels of other pixels are determined using graph cuts [30]. The initial labeling is used to calculate the cost functions in the first round of iteration. After the B-spline grid of each image is updated, the images are transformed and the segmentation labels also updated based on the transformed images. We formulate the data cost as a combination of two terms which, individually, exploit the different characteristics of the perfusion datasets. E_{data} is defined as

$$E_{\text{data}} = w_1 E_{\text{W}} + w_2 E_{\text{Q}}, \tag{7}$$

where E_{Q} calculates the *total quadratic variation* of the dataset, E_{W} calculates the *within-class distance* of each pixel, and w_1 and w_2 are weights that determine the relative

contribution of each term. w_1=0.4 and w_2=1. Two weights are used in order to examine the relative contribution of each term to the results (discussed in "Importance of EQ and EW"). Below we explain each term in greater detail.

Total Quadratic Variation (EQ)

After registration of DCE-MRI we expect the intensity time variation of pixels to be smooth due to motion correction. The *total quadratic variation*, *EQ*, measures the smoothness of the intensity signal by a combination of its first and second order temporal derivatives. The sum of first derivatives over the entire sequence contributes to a smooth signal while the sum of second derivatives favors a piecewise linear signal. During the pre- and post-contrast stages, or for regions without contrast enhancement, the first order derivative is relevant because we expect the intensity of the same tissue to remain constant (i.e., the first order derivative of a constant signal is zero). During the wash-in and wash-out stages, we expect the intensity of the same tissue to increase or decrease approximately with a constant rate. In an unregistered image sequence, the derivatives during the wash-in and wash-out stages will alternate between positive and negative values. However, as the image sequence is registered, the intensity changes are gradual and sum of derivative values minimal. We also use the second-order derivative of the intensity vector to encourage a piecewise linear intensity signal, as the second derivative of a linear signal is zero. Except for position of peak enhancement, the second derivative is zero at other time points.

Let $I(t)$ denote the image at the tth time point (or the tth frame in the dynamic sequence. The intensity of its nth pixel is given by $In(t)$, where $t = 1; \cdots; T$. Thus E_Q is given by

$$E_Q = \sum_{n=1}^{N} \sum_{t=0}^{T} \left(I_n^{'}(t) + I_n^{''}(t) \right) \tag{8}$$

Here, I' and I'' are, respectively, the first and second derivatives of the intensity signal, and N is the total number of pixels in each image. Note that both the first and second derivatives are used for the entire image sequence.

Within-Class Distance (E_W)

The within-class distance, E_W, integrates segmentation information into the JSGR framework. E_W ensures that pixels of the same class have similar intensity time profiles. In other words, pixels with the same label are made similar to a representative signal from that class. We use the mean intensity vector for representing each class of pixels. For a pixel i with known label l, its intensity vector provides greater information about the class labels. The within-class distance is calculated as the difference with respect to the mean intensity vector of class l ($\bar{I}_l$), and is given by

$$E_W = \sum_i \left\| I_i - \bar{I}_l \right\| \tag{9}$$

$\|.\|$ denotes $L2$ norm of the vectors. Note that if pixel i belongs to class l (as determined from the current labels) then its difference only w.r.t., the mean intensity vector of class l is calculated. If a pixel has been correctly labeled as LV blood pool (or RV), then the residual error from the mean of LV blood pool (or RV) will be low. On the other hand, if the labeling is wrong, then the corresponding error is high. Initially, due to many unregistered pixels the mean intensity vector may not be a very accurate representation of the class. But after every iteration, the mean intensity vector becomes smooth with the update of segmentation labels and starts to truly represent the particular class. In an iterative method these constraints (E_Q and E_W) ensure that the labels converge correctly (discussed in "Convergence of Labels"). Although the blood pool shows a lot of intensity change, a combination of E_Q and E_W overcomes these effects.

Here, we need a representative intensity vector for a particular class. Although the mean vector is not necessarily the best representative vector at the beginning of registration, it is the best choice for a balance between registration accuracy and computational complexity. PCA could be a more accurate choice for the representative vector, but significantly increases computation complexity. We observe that with increasing number of iterations the mean vector does converge to the representative vector of the class.

Optimization

The graph cut algorithm of Ref. [30] is used to get the segmentation labels. Pixels are represented as nodes in a graph G which consists of a set of directed edges that connect two nodes. The edge weight between two neighboring nodes is the smoothness term while the data penalty term is the edge weight for links between nodes and label nodes (terminal nodes). Edges between pixel nodes are called n-links while edges between pixel nodes and terminal nodes are called t-links. The optimum labeling set is obtained by severing the edge links in such a manner that the cost of the cut is minimum. Cost of a cut is the sum of the weights of the severed edges. For l labels, there are l terminal nodes (for our experiments l=4, i.e., LV, RV, myocardium, and background). The number of nodes equals N_p, the number of pixels. The optimal labeling is obtained via a series of expansion moves. Details of graph construction and optimization can be found in [30].

B-splines are used to optimize the cost function in Eq. 1. A uniform grid is initialized for all images. The first image's

grid coefficients are updated based on the present value of the energy function. The image is transformed, and the segmentation labels are immediately updated. The next image's grid coefficients are then updated followed by image transformation and update of segmentation labels. This is repeated for all images of the sequence corresponding to different time points. This constitutes one iteration for JSGR. The updated segmentation labels are used as the starting point for the next iteration. We repeat the process until the cost function does not decrease further. A brief discussion is presented in "Convergence of Labels."

The advantage of dynamic update of segmentation labels is that it reflects the value of the energy function of the updated image sequence based on the latest transformations. Thus, if the energy function value does not change while updating the grid coefficients of three consecutive images, then JSGR is terminated and the final segmentation labels obtained using graph cuts. This leads to accurate registration and lesser iterations than updating the segmentation labels at the end of each iteration. To register the images to a common image space, the average deformation of a pixel over all time points is made zero by making the sum of B-spline coefficients for the corresponding grid points of all images to be zero. The method is summarized in Algorithm 1.

Experiments and Results

Cardiac images were acquired on a 1.5 T Siemens Sonata MR scanner following bolus injection of Gd-DTPA contrast agent. There are 12 sequences acquired from 12 patients in whom it was important to look at myocardial perfusion. Each sequence comprised of 60 frames with a total of 720 frames. The datasets were acquired with electrocardiographic gating such that the images were acquired during the same phase of the cardiac cycle. This minimized cardiac motion but some deformations were still observed due to patient breathing. The pixel spacing ranges from $(1.5\times 1.5)-(2.8\times 2.8)\text{mm}^2$. The acquired images were from the same mid-cavity slice and their dimensions was between $84-92\times 93-102$ pixels. The images were corrected for rotation and translation motion before segmentation. The initial labeling was obtained using graph cuts [30]. The B-spline grid for JSGR was of size 10×10 with the spacing between grid control points varying from $7-9$ pixels.

Registration Results

We compare the registration from JSGR with [13] Met 1 and [28] Met 2. Met 1 does not define an explicit reference image and uses the entropy of the pixel stack for the cost function. It was implemented in ITK using line search with gradient descent. The initial coarse B-spline grid had pixel spacing of 22 pixels, and two subsequent finer resolutions had pixel spacing of 16 and 9 pixels. This multiresolution scheme avoids local minima. Stack entropy was the cost function. Apart from the grid size, there was no other parameters that need to be specified.

Met 2 uses pairwise registration with qualitative mutual information [29] as the similarity measure. It has been used before for registering contrast enhanced cardiac MRI. A multiresolution B-spline grid was used with control point spacing of 22, 16, and 9 pixels at three resolutions. Our algorithm converged after six iterations. The threshold cost difference above which registration continues is set at 0.1. We present qualitative and quantitative results of our method (JSGR) in terms of contour distance and segmentation accuracy. For Met 1, the number of iterations required for convergence was five while on an average, Met 2 needed four iterations for convergence while registering a pair of images. The reported results are after convergence of all algorithms.

The outline of the LV, myocardoium, and RV are manually identified by expert observers in the original image sequence. These contours are denoted as C_{org}. The transformation of each image is used to map the contours to the registered image space. Let these contours be denoted as C_{trans}. Since it is practically impossible to have ground truth value for elastic registration, we take a different approach to quantify registration accuracy. In a perfusion image

Table 1 Summary of registration and segmentation performance on cardiac perfusion datasets

	Registration results (CD in mm)				Segmentation results		
	Before registration	After registration			DM (%)		
		JSGR	Met 1	Met 2	JSGR	Met 1	Met 2
Myocardium	2.2±1.2	0.7±0.1	1.0±0.2	1.1±0.1	93.6±1.7	90.1±1.3	90.6±1.6
LV	2.6±1.1	0.5±0.2	0.9±0.3	1.0±0.1	93.1±1.5	89.4±1.2	88.8±1.8
RV	2.4±1.2	0.7±0.1	1.1±0.2	1.3±0.2	92.4±1.4	88.8±1.3	87.2±1.9

The values indicate average and standard deviations for all datasets

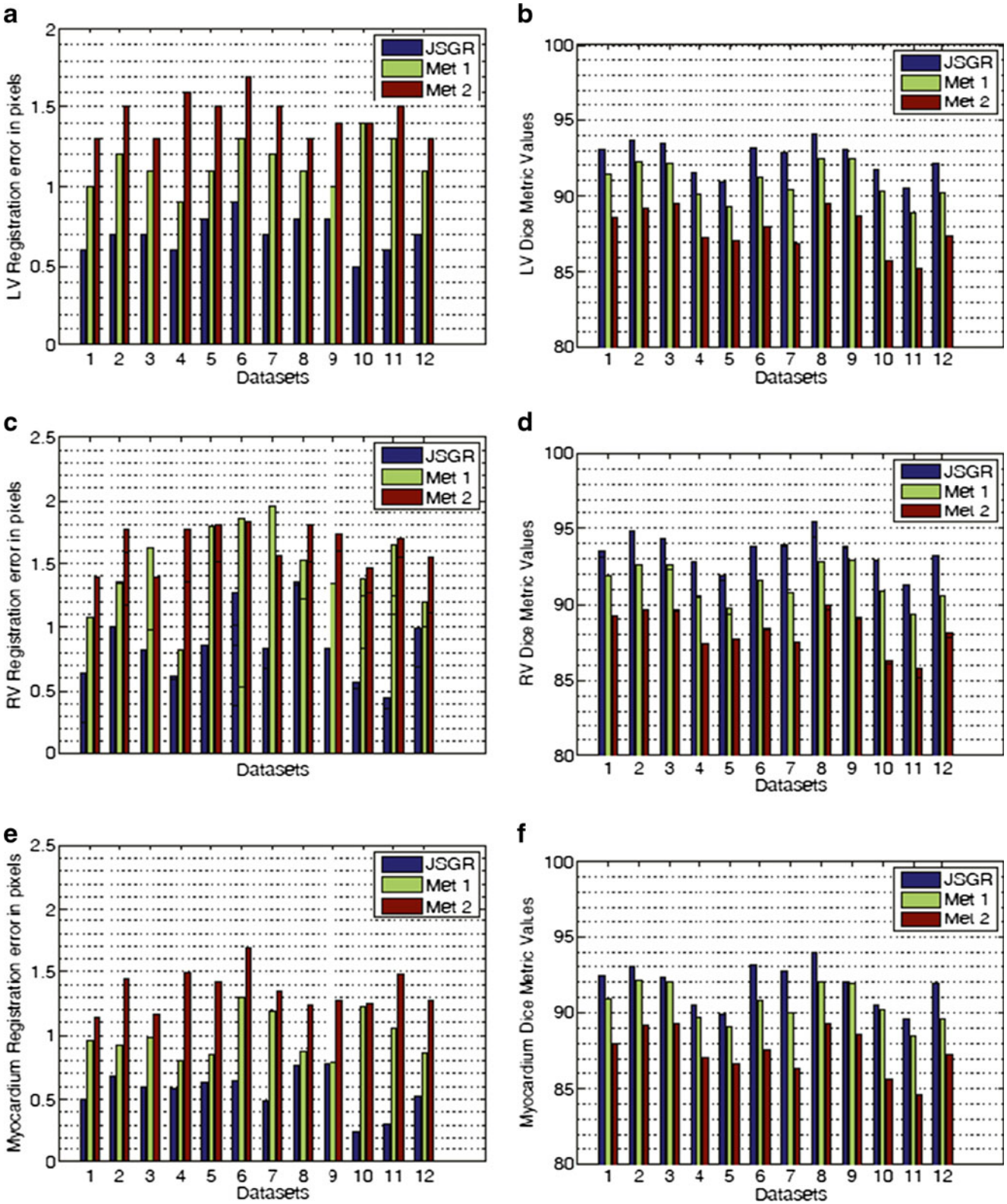

Fig. 1 Registration and segmentation results for LV, RV, and myocardium for each of the 12 datasets **a**, **c**, **e** CD values in pixels; **b**, **d**, **f** DM values in percent

sequence, the shape of the heart should be same for ideal registration of all images. We fix one image as reference (usually the first image of the sequence) and calculate the contour distance between the LV (or any object of interest)

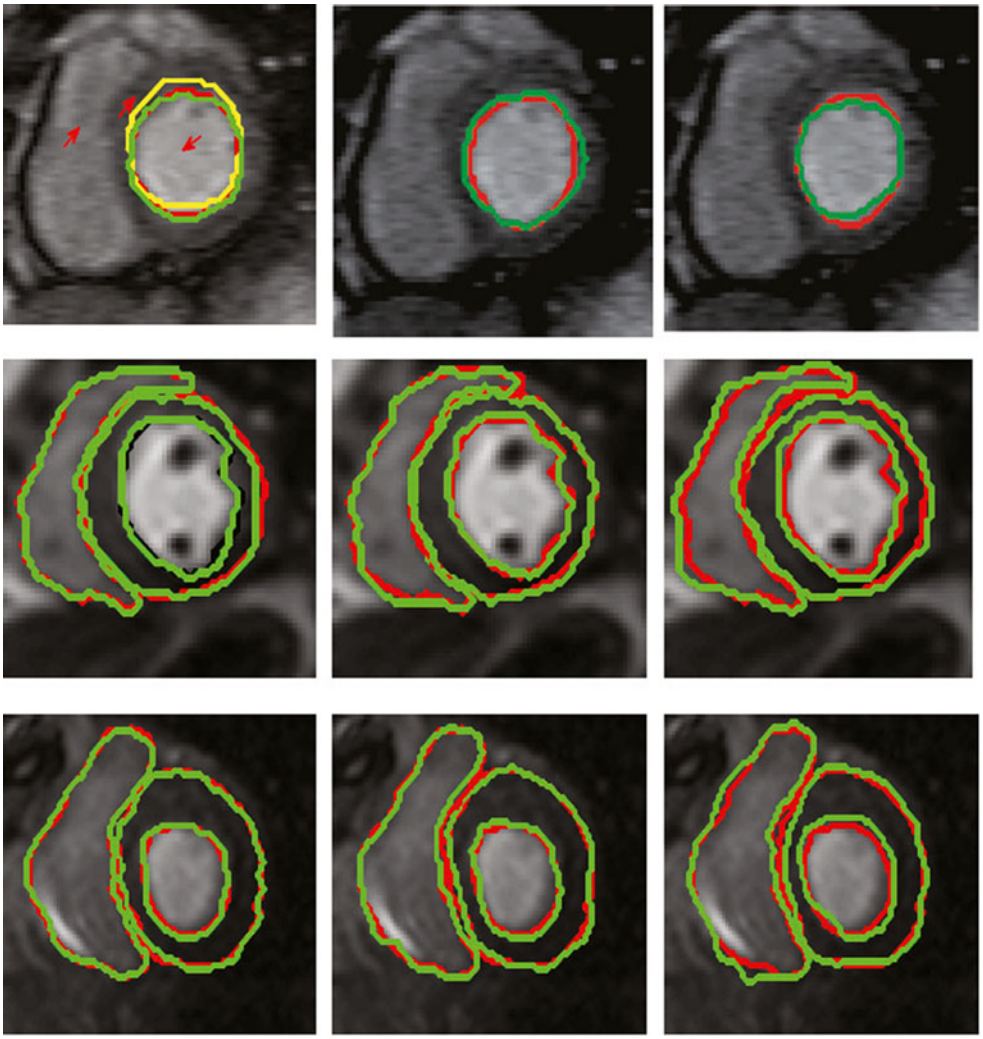

Fig. 2 Contours of segmented regions overlaid on images from the dataset. First column shows results for JSGR, second column for Met 1, and third column for Met 2. Each row corresponds to a different dataset. *Red* contours show the manual segmentations while the *green* contours show the automatic segmentations. The initial segmentation for JSGR is shown in *yellow*. First row shows only one set of results for the clarity

in all other images and the reference image. At the end of our JSGR method, not only has the sequence been registered but the segmentation labels of each frame have also been determined. Thus we just need to find the contour distance between regions with the same labels. By contour we refer to the outline of the segmented region. Note that the same procedure as above can be used to calculated the contour distance before registration. The contour distance is calculated using the following steps

1. Let the reference contour for the registered image sequence be denoted as $C_{\text{trans}}^{\text{ref}}$.
2. Let the i^{th} point on C_{trans} be denoted by $C_{\text{trans}}(i)$ and the j^{th} point on $C_{\text{trans}}^{\text{ref}}$ be denoted by $C_{\text{trans}}^{\text{ref}}(j)$.

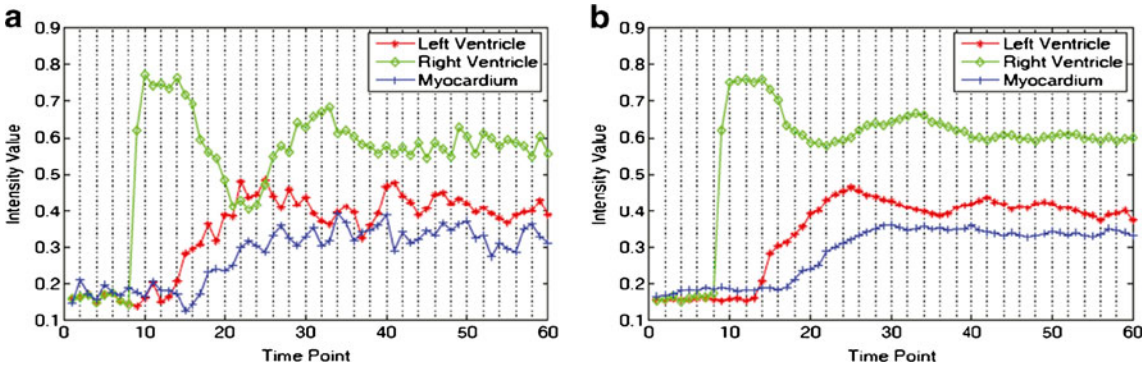

Fig. 3 Intensity change with time for pixels on epicardium, RV, and endocardium; **a** before groupwise registration; **b** after groupwise registration

3. For every $C_{trans}(i)$ find the point on $C^{ref}_{trans}(j)$such that the distance between $C_{trans}(i)$ and $C^{ref}_{trans}(j)$ is minimum.

$$d(i,j) = \min j \left\| C_{trans}(i) - C^{ref}_{trans}(j) \right\| \quad (10)$$

4. For each point $C_{trans}(i)$ the corresponding $d(i,j)$ is calculated. The contour distance (CD) is the average distance and is defined as

$$CD = \frac{\sum_i d(i,j)}{n} \quad (11)$$

n is the number of points on the contour.

The average *CD* values for 12 datasets are given in Table 1. Lower CD implies better performance. We observe that JSGR has the lowest CD values while Met 2 has the highest values. Met 2 registers all images to a fixed reference image. Intensity change is common between two images of the dataset. Consequently, registration is prone to error while using a simple NMI based similarity measure. Groupwise registration has the advantage that information from the entire sequence can be exploited for registration which is particularly important for DCE-MRI. Between Met 1 and JSGR, the latter performs better because it combines segmentation and registration information in the cost function. Subsequently, the final CD values for JSGR are lower than Met 1. In Fig. 1a, c, e, we show the average CD values for the LV,RV and myocardium for each of the 12 datasets.

The time taken for registering one full dataset (including 60 images) is 1 h and 33 min using JSGR, 58 min using Met 1 and 50 min using Met 2. JSGR was implemented using MATLAB 7.5 on a PC having a Pentium 4, 3 GhZ processor. Met 1 was implemented by the authors using ITK and thus has low execution time. Met 2 was also implemented in MATLAB for pairwise registration. Since JSGR is a joint segmentation and groupwise registration method, it takes more than twice the time compared with Met 2.

Segmentation Results

Segmentation accuracy is calculated based on dice metric (DM) values between manual segmentation and automatic segmentation for different methods. After registration is complete for each method the segmentation labels are obtained by applying graph cuts on the intensity vectors. The segmentation labels for corresponding pixels on different slices will be the same. For JSGR, the labels are already obtained after the registration process. The average DM values for the LV over all 12 datasets are shown in Table 1. Figure 1b, d, f shows the average DM values of the LV, RV, and myocardium for each dataset. The DM values are highest for JSGR thus indicating maximum accuracy amongst all methods.

Figure 2 shows the segmented contours for LV, RV, and myocardium overlaid on a representative image of the database. The representative image is chosen such that the blood pool is visible without any ambiguities. The manual segmentations are shown in red while the automatic segmentations are shown in green. JSGR requires initial segmentation which are shown in the first row in yellow. The initial segmentation is a result of applying graphcut to the unregistered image sequence and using the intensity vector of each pixel. Note that the automatic segmentations are the average contours over all frames of the sequence and the manual segmentations are also the average of the manually drawn contours. Since the segmentation labels are calculated from the intensity vectors, the labels will be the same for corresponding pixels in all frames. The first column in Fig. 2 shows results for JSGR, second column shows results for Met 1, and the third column shows results for Met 2. Each row shows results for different datasets. Again we observe that JSGR shows the best agreement with manual segmentations due to the combination of registration and segmentation information. The other two methods being solely focused on registration perform inadequately for cardiac perfusion images. The accuracy measures highlight the importance of integrating registration and segmentation information. This combination is particularly important when images have low contrast, and also when segmentation information is available to be exploited for registration.

Figure 3 shows the intensity variations with time for pixels on the LV blood pool, RV, and myocardium. Since edge pixels are chosen the intensity change before registration (Fig. 3a) is noisy. After registration using JSGR, the intensity variation is smoother (Fig. 3b) and highlights the success of our method for time varying data.

Importance of E_Q and E_W

It is important to look at the contribution of E_Q and E_W to the overall registration procedure. E_Q can be termed as the registration energy while E_W is the segmentation energy. We examine the improvement brought about by E_W to registration accuracy, and also the improvement in segmentation accuracy due to E_Q. We vary w_1 (Eq. 7) from 0 to 5 (keeping w_2=1 fixed) and calculate the registration and segmentation accuracy values for the LV (shown in Table 2). It is observed

Table 2 Average registration and segmentation accuracy of LV with change in w_1

w_1	0	0.1	0.2	0.4	1	1.5	2	2.5	3	4	5
CD (pix)	1	0.9	0.7	0.6	0.6	0.7	0.7	0.8	0.9	1.2	1.5
DM (%)	89.3	90.6	91.9	93.8	93.7	93.1	92.9	91	89.9	89.3	88.8

Table 3 Average registration and segmentation accuracy of LV with change in w_2

w_2	0	0.1	0.2	0.4	1	1.5	2	2.5	3	4	5
CD (pix)	1.4	1.2	1	0.8	0.6	0.6	0.7	0.75	0.8	1.1	1.4
DM (%)	88.2	89.5	90.4	91.7	93.8	93.5	92.7	92.1	91.8	89.2	88.3

that as w_1 increases from 0, the registration and segmentation performance both improve. In fact when $w_1 \leq 0.4$, the registration accuracy is less than or comparable to Met 1 but improves with greater contribution of E_W. However, if $w_1 > 2$ the registration accuracy starts to degrade (evident from higher CD) and the DM values also decrease due to unbalancing of each terms contribution.

Similarly, when we increase the value of w_2 (with $w_1=0.4$) from 0 to 5 we observe low DM values when $w_2 < 1$. The best segmentation accuracy is obtained for $1 \leq w_2 \leq 3$ However, when $w_2 > 3$ the DM starts to decrease and the CD also increases. Table 3 shows the change in CD and DM with change in w_2. The best results are obtained for $w_2=1$ and $w_1=0.4$.

Convergence of Labels

The iterative process converges when there is no further decrease of the total energy. Figure 4a shows the change in cost with each iteration of a particular dataset using all three registration methods. Iteration 1 indicates the cost function value before registration. This value is generally large because all the images are misaligned. After iterations 2 and 3, the change in cost function value is larger than subsequent iterations because of the large motion correction initially. This is particularly true for the iteration 2. For subsequent iterations, the decrease in cost function is not so large because the registration only corrects small local deformations.

In another set of experiments, we deliberately assigned erroneous segmentation labels in the initial segmentation, and applied JSGR. Figure 4b shows the change in cost function with each iteration for datasets 5 and 7. In this particular scenario, the cost function increases after the first iteration because of the inaccurate initial labeling. For subsequent iterations, the cost decreases and converges to nearly constant value. Plots are shown only for JSGR as the other methods are not dependent upon the initial labeling.

Conclusions

We have proposed a novel method for joint segmentation and groupwise registration of cardiac perfusion images using temporal dynamics. By maximizing the smoothness of the temporal intensity signal, our method uses available temporal information from the entire image sequence. This helps to overcome the effects of intensity change. Segmentation information is incorporated by minimizing the error between a pixel's intensity vector and mean intensity vector of the same class. Compared with manual segmentations, our method gives higher segmentation accuracy and lower registration error than other methods. Our being a *joint segmentation and groupwise registration* approach, both registration and segmentation performance is better than conventional methods. This is because of two factors: (1) exploiting temporal information from DCEMRI sequence in a *groupwise registration* framework and (2) use of mutually complementary registration and segmentation information while most other methods solve registration and segmentation separately. Our method has the potential to be used for other data types having time varying characteristics, and in the future, we aim to use our method on other dynamic datasets.

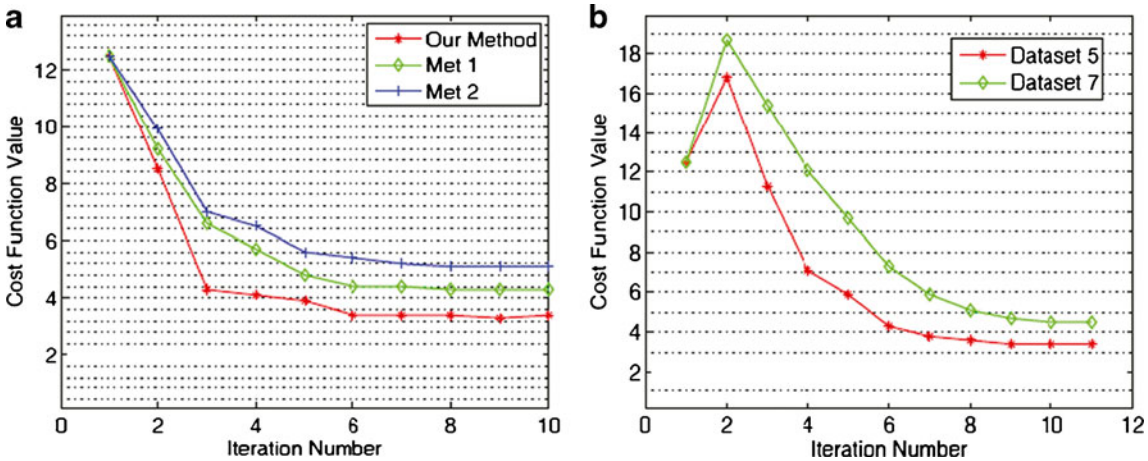

Fig. 4 Values of different cost function with different iterations: **a** for all three methods and **b** for JSGR on datasets 5 and 7 for erroneous initial segmentation

Acknowledgment The author would like to thank Dr. Ying Sun of National University of Singapore for providing the cardiac datasets.

References

1. Mahapatra D, Sun Y: Mrf based intensity invariant elastic registration of cardiac perfusion images using saliency information. IEEE Trans Biomed Eng 58(4):991–1000, 2011
2. Song T, Lee VS, Rusinek H, Kaur M, Laine AF: Automatic 4-D registration in dynamic MR renography based on over-complete dyadic wavelet and Fourier transforms. Med Image Comput Comput Assist Interv 2:205–213, 2005
3. Lee VS, Rusinek H, Bokacheva L, Huang A, Oesingmann N, Chen Q, et al: Real function measurements from mr renography and a multicompartmental model. Am J Physiol Renal Physiol 292:1548–1559, 2007
4. Boykov Y, Lee VS, Rusinek H, Bansal R: Segmentation of dynamic N-D data sets via graph cuts using Markov models. In: Proc. 7th International Conference on Medical Image Computing and Computer Assisted Intervention, 2001, pp 1058–1066
5. Hackstein N, Heckrodt J, Rau WS: Measurement of single-kidney glomerular filtration rate using a contrast enhanced dynamic gradient-echo sequence and the Rutland–Patlak plot technique. J Magn Reson Imaging 18:714–725, 2003
6. Joshi S, Davis B, Jomier M, Gerig G: Unbiased diffeomorphic atlas construction for computational anatomy. Neuroimage 23:151–160, 2004
7. Shattuck DW, Mirza M, Adishetiyo V, Hojatkashaniand C, Salamon G, Narr KL, Poldrack RA, Bilder RM, Toga AW: Construction of a 3D probabilistic atlas for human cortical structures. Neuroimage 30 (3):1064–1080, 2008
8. Lepore N, Brun C, Pennec X, Chou YY, Lopez OL, Aizenstein HJ, Becker JT, Toga AW, Thompson PM: Mean template for tensor based morphometry using deformation tensors. Med Image Comput Comput Assist Interv 2:826–833, 2007
9. Zollei L, Stevens A, Huber K, Kakunoori S, Fischl B: Improved tractography alignment using combined volumetric and surface registration. Neuroimage 51(1):206–213, 2010
10. Yang J, Shen D, Davatzikos C, Verma R: Diffusion tensor image registration using tensor geometry and orientation features. Med Image Comput Comput Assist Interv 11:905–913, 2008
11. Barmpoutis A, Vemuri BC: Groupwise registration and atlas construction of 4th order tensor fields using the R^+ Riemannian metric. Med Image Comput Comput Assist Interv 5761:640–647, 2009
12. Zhang BB, Avants PA, Yushkevich JH, Woo S, Wang LF, McCluskey LB, Elman ER, Gee JC: High dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis. IEEE Trans Med Imag 26(11):1585–1597, 2007
13. Balci SK, Golland P, Shenton M, Wells WM: Free-form B-spline deformation model for groupwise registration. Med Image Comput Comput Assist Interv 10:23–30, 2007
14. Learned-Miller EG: Data driven image models through continuous joint alignment. IEEE Trans Patt Anal Mach Intell 28(2):236–250, 2006
15. Baloch S, Verma R, Davatzikos C: An anatomical equivalence class based joint transformation-residual descriptor for morphological analysis. Inf Process Med Imaging 20:594–606, 2007
16. Jia H, Yap P-T, Wu G, Wang Q, Shen D: Intermediate templates guided groupwise registration of diffusion tensor images. Neuroimage 54(2):928–939, 2011
17. Hamm J, Davatzikos C, Verma R: Efficient large deformation registration via geodesics and learned manifold of images. Med Image Comput Comput Assist Interv 12:680–687, 2009
18. Tang S, Fan Y, Shen D: RABBIT: rapid alignment of brains by building intermediate templates. Neuroimage 47(4):1277–1287, 2009
19. Zhang L, Chefd'hotel C, Bousquet G: Group-wise motion correction of brain perfusion images. In: 2010 IEEE International Symposium on Biomedical Imaging From Nano to Macro. Siemens Corp. Res., Princeton, NJ, USA, 2010, pp 832–835
20. Metz CT, Klein S, Schaap M, van Walsum T, Niessen WJ: Non rigid registration of dynamic medical imaging data nD+t B-splines and a groupwise optimization approach. Med Image Anal 15 (2):238–249, 2011
21. Yezzi A, Zollei L, Kapur T: A variational framework for joint segmentation and registration. In: IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. IEEE Computer Society Washington, DC, USA, 2001, pp 44–51
22. An J, Chen Y, Huang F, Wilson D, Geiser E: A variational PDE based level set method for a simultaneous segmentation and non-rigid registration. Med Image Comput Comput Assist Interv 8:286–293, 2005
23. Pohl KM, Fisher J, Grimson WEL, Kikinis R, Wells WM: A Bayesian model for joint segmentation and registration. Neuroimage 31(1):228–239, 2006
24. Mahapatra D, Sun Y: Joint registration and segmentation of dynamic cardiac perfusion images using MRFs. Med Image Comput Comput Assist Interv 13:493–501, 2010
25. Mahapatra D, Sun Y: Integrating segmentation information for improved MRF based elastic image registration. IEEE Trans Imag Proc 21(1):170–183, 2012
26. Wu G, Jia H, Wang Q, Shen D: Groupwise registration with sharp mean. Med Image Comput Comput Assist Interv 13:570–577, 2010
27. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ: Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imag 18:712–721, 1999
28. Mahapatra D, Sun Y: Registration of dynamic renal MR images using neurobiological model of saliency. In: 2010 IEEE International Symposium on Biomedical Imaging From Nano to Macro. Siemens Corp. Res., Princeton, NJ, USA, 2008, pp 1119–1122
29. Sun Y, Jolly M-P, Moura JMF: Contrast-invariant registration of cardiac and renal MR perfusion images. In: Proc. 7th International Conference on Medical Image Computing and Computer Assisted Intervention. France, UK, 2004, pp 903–910
30. Boykov Y, Funka-Lea G: Graph cuts and efficient N-D image segmentation. Intl J Comp Vis 70(2):109–131, 2006